Received: 1 January 2019 | Revised: 25 April 2019 | Accepted: 16 May 2019 | Published on: 20 June 2019

DOI: 10.1002/qj.3568

RESEARCH ARTICLE

Quarterly Journal of the Royal Meteorological Society RMetS

Application of the Warn-on-Forecast system for flash-flood-producing heavy convective rainfall events

Nusrat Yussouf[1,2,3] | Kent H. Knopfmeier[1,2]

[1]Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

[2] National Oceanic and Atmospheric Administration, National Severe Storms Laboratory, Norman, Oklahoma

[3]School of Meteorology, University of Oklahoma, Norman, Oklahoma

Correspondence
Nusrat Yussouf, National Oceanic and Atmospheric Administration, National Severe Storms Laboratory, National Weather Center, 120 David L. Boren Boulevard, Norman, OK 73072.
Email: nusrat.yussouf@noaa.gov

Funding information
NOAA-University of Oklahoma Cooperative Agreement #NA11OAR4320072, U.S. Department of Commerce.

Abstract

The vision of the National Oceanic and Atmospheric Administration's (NOAA) Warn-on-Forecast (WoF) research and development project is to provide very short-term probabilistic model guidance products that will aid the ability of National Weather Service forecasters to issue probabilistic warnings of severe convective hazards with higher accuracy and longer lead times. The experimental Warn-on-Forecast System (WoFS), which is under development, is a frequently cycled, regional, convective-scale, on-demand, ensemble data assimilation and prediction system. The system assimilates thunderstorm observations to forecast the life cycle and associated hazards of individual convective storms. Most of the initial research effort, since the beginning of the WoF project in 2009, has been focused on tornadic events, which constitute one of the most violent and difficult to predict weather threats. However, the system emerging from this project has significant potential to help with prediction of other hazards as well: in particular, flash-flood-producing, convectively driven intense rainfall events. This study focuses on the application of WoFS in 0–6-hr probabilistic rainfall forecasts of several flash-flood-producing heavy rainfall events during the spring and summer of 2015 and 2016. Results indicate that WoFS successfully initializes the convective storms in the ensemble. Visual inspections and probabilistic verification metrics show that WoFS predicts the intense rainfall that often leads to flash flooding at the correct location with higher accuracy in areal coverage and amount during the 0–3 hr forecast period than during the later 3–6 hr period. Overall results indicate that the frequently updated 0–6-hr ensemble forecast from the WoFS has the potential to highlight areas where intense rainfall can result in flash flooding and increase near-term situational awareness of flash flood threats.

KEYWORDS
convective-scale numerical modeling, ensemble data assimilation, ensemble forecasting, flash flood, heavy rainfall, Warn-on-Forecast

1 | INTRODUCTION

The Warn-on-Forecast (WoF) project (Stensrud *et al.*, 2009, 2013) at the National Oceanic and Atmospheric Administration's (NOAA) National Severe Storms Laboratory (NSSL) aims to enable a new paradigm for National Weather Service (NWS) severe weather watch-to-warning operations, where convective-scale probabilistic numerical weather prediction (NWP) model guidance is the key resource. The outcome of NSSL's several years of research and development efforts (Dawson *et al.*, 2012; Yussouf *et al.*, 2013a, 2013b, 2015; Wheatley *et al.*, 2014; 2015; Jones *et al.*, 2016; Skinner *et al.*, 2016) since the beginning of the WoF initiative in 2009 is an experimental Warn-on-Forecast System (WoFS). The system is a rapidly updating, regional, convective-scale, on-demand, ensemble data assimilation (DA) and prediction system that provides very short-term probabilistic model guidance for tornadoes, heavy rainfall, large hail, damaging straight-line winds and other convective hazards. The WoFS developers are working on continuous improvements and enhancements of the system, based on findings from rigorous quantitative and qualitative case study evaluations and real-time testing during peak severe weather season in the United States.

The initial focus of the WoFS is to improve ∼1–2-hr probabilistic forecasts of low- and mid-level rotations of individual convective storms (Yussouf *et al.*, 2013a, 2013b, 2015; Wheatley *et al.*, 2015). While tornadoes are one of the most violent severe weather hazards, floods and flash floods are a major cause of weather-related deaths and property damage in the United States (www.nws.noaa.gov/om/hazstats.shtml) and around the world (e.g., https://public.wmo.int/en/media/news/new-report-highlights-economic-cost-of-disasters). For instance, in 2015, a combination of river and flash flooding was responsible for the largest number of weather-related fatalities for the year, almost double that of tornadoes and high-temperature-related deaths combined. Therefore, it is crucial that forecasters have accurate model prediction of intense rainfall events that have the potential to produce flash flooding. As such, accurate short-term probabilistic model forecasts of flash-flood-producing heavy convective rainfall events are essential and align with the objectives of NOAA's WoF project.

Recent work by Yussouf *et al.* (2016) demonstrates the potential of the WoFS to produce skillful short-term extreme rainfall forecasts. This study of the May 31, 2013 Oklahoma, USA tornado and flash flood event shows that the system predicts not only the high low-level mesocyclones associated with the El Reno tornado but also short-term intense rainfall that matches the observed rainfall well in terms of location and amount. The ensemble-derived probabilistic rainfall guidance from assimilating operational Weather Surveillance Radar-1988 Doppler (WSR-88D) reflectivity and radial velocity observations is more skillful than the rainfall forecasts from an experiment with no radar DA. Heavy rainfall associated with deep moist convective systems often occurs on small spatial and temporal scales and numerical prediction of the location, timing, and intensity of rainfall is challenging (Sun *et al.*, 2014; Schumacher, 2017 and references therein). Even though the convective-scale numerical models are constrained by rapid error growth, the forecast length of the WoFS is extended out to 6 hr to evaluate the potential utility of the system in predicting a flash flood threat, which is defined as "the flooding that begins within 6 hr, and often within 3 hr, of the heavy rainfall or other cause" (https://www.weather.gov/phi/FlashFloodingDefinition).

To explore the robustness of WoFS in forecasting short-term intense convective rainfall, the current study simulates multiple retrospective heavy rainfall and flash flood events from the 2015 and 2016 warm season. The WSR-88D reflectivity, radial velocity, and all other available conventional observations are assimilated into the frequently cycled, 36-member WoFS ensemble at 3-km horizontal grid spacing. A continuum of 0–6-hr ensemble forecasts is initialized from the frequently adjusted ensemble analyses as new storm observations become available. The goal is to assess how accurately the system can predict the location, timing, and intensity of heavy rainfall, which in turn can enhance the forecast of flash floods with longer lead time.

An overview of the heavy rainfall events that are simulated retrospectively in this study is given in section 2. The experimental WoFS configuration is described in section 3. Section 4 discusses the DA performance and the ensemble forecasts. A summary and concluding remarks are found in section 5.

2 | SUMMARY OF THE EVENTS

The NOAA Weather Prediction Center's (WPC) Day 1 excessive rainfall outlook (ERO) for five significant flash-flood-producing heavy rainfall events (Figure 1) simulated using the WoFS ranges from slight to high risk (Table 1). Convection started during the late afternoon hours and continued into the evening hours and overnight to the next day. The timing and location of the first flash flood report for events of interest are also listed in Table 1. The convective-scale regional 3-km domains used in each experiment are shown in Figure 2, with the local storm reports obtained from the National Center for Environmental Information's (NCEI) *Storm Data* publication. A brief overview of the five heavy rainfall events used in this study is given below.

2.1 | April 26–27, 2015 north–central Texas event

A fairly classic severe weather setup for the southern Plains was present on April 26, 2015. Convergence along the dryline promoted convection initiation around 1800 UTC.

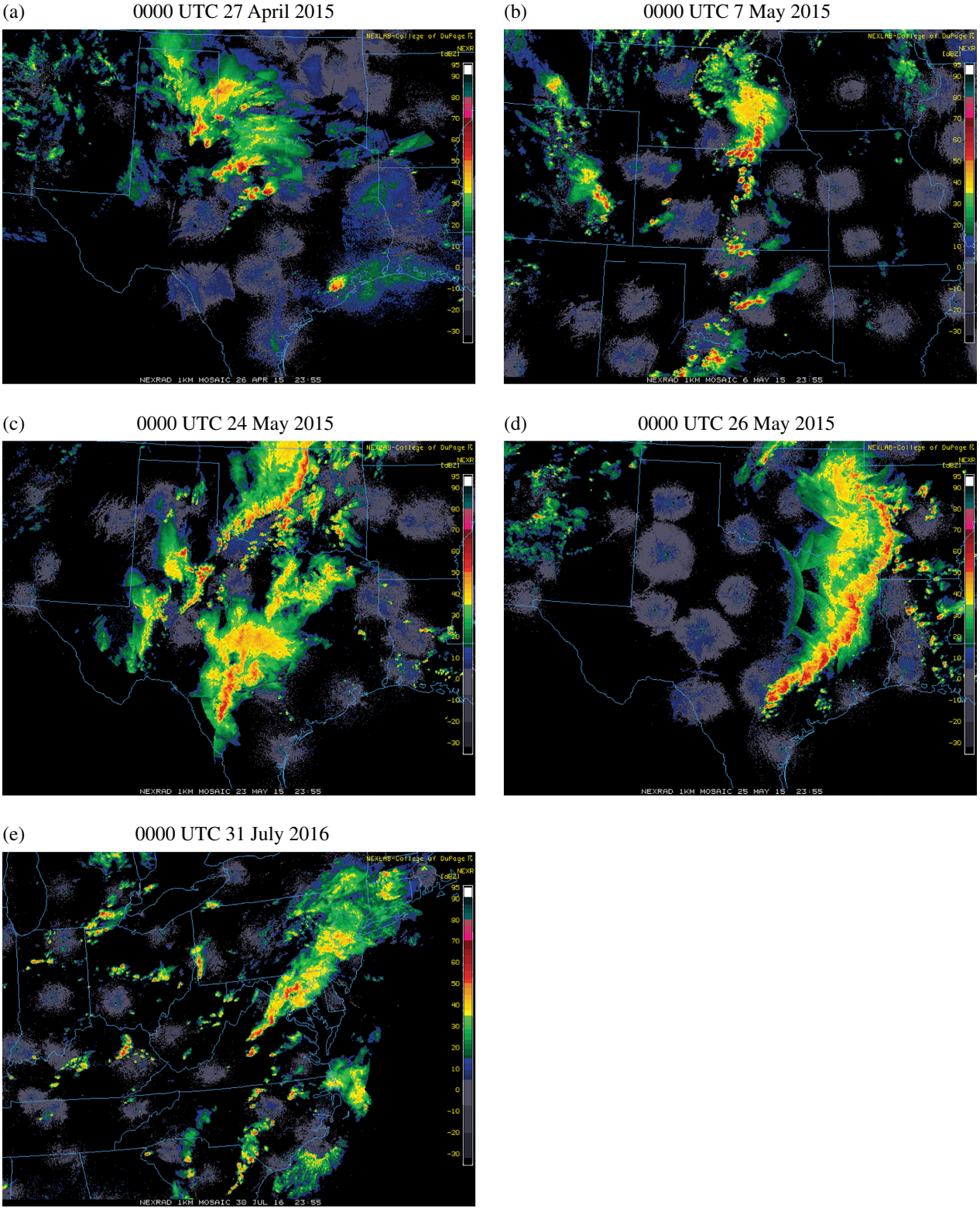

FIGURE 1 Composite reflectivity (courtesy of http://www2.mmm.ucar.edu/imagearchive/) valid at 0000 UTC on (a) April 27, 2015, (b) May 7, 2015, (c) May 24, 2015, (d) May 27, 2015, and (e) July 31, 2016 [Colour figure can be viewed at wileyonlinelibrary.com].

The combination of instability with a strong deep layer and low-level shear favored the development of supercells from the incipient convection. Numerous reports of large hail and tornadoes occurred over north–central Texas during the late afternoon hours (Figure 2a). A pair of supercell thunderstorms stalled and eventually merged over Johnson and Ellis counties during the late evening hours (Figure 1a), producing heavy rainfall and causing significant flash flooding. Several roads were flooded or washed out across the region.

TABLE 1 The list of flash-flood-producing heavy convective rainfall events used in this study

Date	Geographic location	WPC's day[a] excessive rainfall outlook	First flash flood report[a]1 (Date, UTC, County)
April 26–27, 2015	North–central Texas	Moderate	20150427, 03:57 Johnson
May 6–7, 2015	Central Oklahoma	Slight	20150506, 23:13 McClain
May 23–24, 2015	South–central Texas	Moderate to high	20150523, 21:40 Kendall
May 25–26, 2015	Southeast, Texas	Slight to moderate	20150525, 23:52 Brazos
July 30–31, 2016	Maryland	Slight	20160731, 00:18 Montgomery

[a]https://mesonet.agron.iastate.edu/request/gis/lsrs.phtml

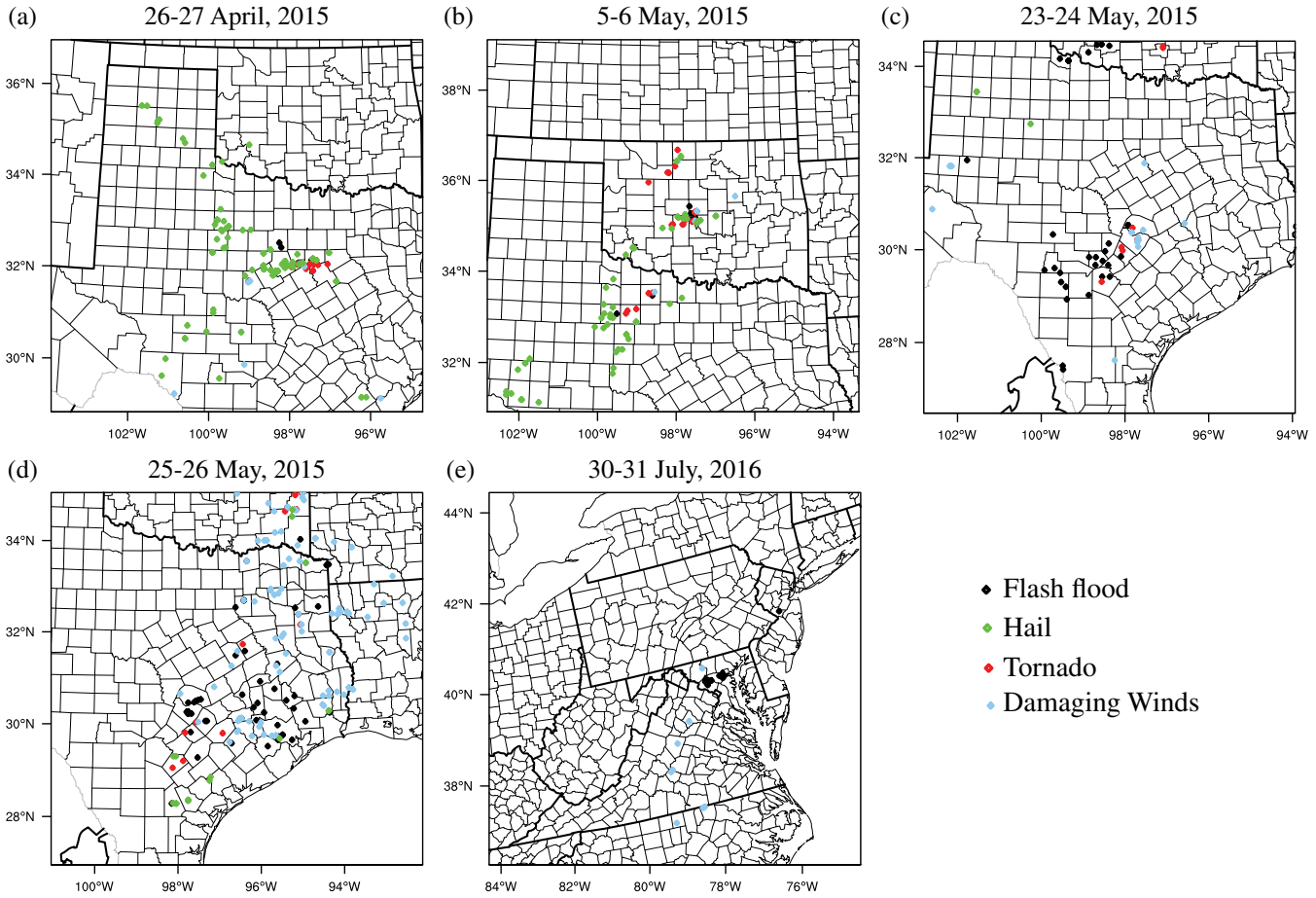

FIGURE 2 The convective-scale domains for the five case studies used in the DA and forecast experiments. Overlaid are the local storm reports obtained from NCEI's Storm Data publication valid between 2100 UTC and 0600 UTC the next day. Details are shown in the legends [Colour figure can be viewed at wileyonlinelibrary.com].

2.2 | May 6–7, 2015 central Oklahoma event

Convective potential was reasonably high over central Oklahoma in the afternoon of May 6, 2015 with ~ 1,500 J/kg of MLCAPE and 40 knots of deep-layer wind shear around 1900 UTC. Supercells began to form across southwestern Oklahoma at ~ 2000 UTC. These storms moved slowly northeast and began to backbuild over the Oklahoma City metropolitan area (Figure 1b), producing anywhere from 1 to over 3 inches of rain over the region. Many locations in the area had received from 2 to more than 4 inches of rain the previous day, resulting in flash flooding in multiple locations (Figure 2b). Fifteen tornadoes were also reported across Oklahoma during the afternoon and early evening hours, including two EF-3s and one EF-1 within the Oklahoma City metropolitan area.

2.3 | May 23–24, 2015 south–central Texas event

The combination of instability and deep-layer shear promoted multiple rounds of heavy thunderstorms over the area from

the late afternoon hours of May 23 to the early morning hours of May 24, which led to widespread rainfall of 6–8 inches (Figures 1c and 2c). Significant flash and river flooding within the Blanco River basin resulted in major impacts to life and property across the region.

2.4 | May 25–26, 2015 southeast Texas event

During the evening hours of May 25, an intense MCS approached the region while isolated supercells developed to the southwest. As the leading edge of the MCS convection sagged southward, it merged with the northeastward-moving supercells over the southwest/western portions of the Houston, Texas metropolitan area (Figures 1d and 2d). Convection continually trained over the area and the western portions of the city of Houston experienced over 152.4 mm (6 inches) of rainfall. The addition of this event to the heavy rainfall of the previous days/weeks produced record flooding across Houston, and resulted in the issuance of a flash flood emergency for that area by the National Weather Service at 0452 UTC on May 26, 2015.

2.5 | July 30–31, 2016 Ellicott City, Maryland event

Torrential rainfall occurred during the early to mid-evening hours in the Ellicott City, Maryland area, causing severe flash flooding and destruction to numerous buildings in the historic Old Town portion of the city (Figure 2e). Several convective cells ahead of a larger area of rain followed by a west–east oriented band (Figure 1e) of intense rainfall produced nearly 5.5 inches of rain in 90 min in Ellicott City.

3 | DESCRIPTION OF AN EXPERIMENTAL WARN-ON-FORECAST SYSTEM

The experimental WoFS uses the Advanced Research Weather Research and Forecasting (WRF-ARW version 3.8.1; Skamarock *et al.*, 2008) NWP model and the Community Gridpoint Statistical Interpolation (GSI: DTC, 2017a; Kleist *et al.*, 2009) based ensemble Kalman filter (EnKF: Houtekamer *et al.*, 2005; Whitaker *et al.*, 2008; DTC, 2017b) DA system (GSI-EnKF hereafter), which is used as a component within the United States NWS operational DA system (Wang *et al.*, 2013). The configuration of the experimental WoFS used to conduct this study is similar to that used by Yussouf *et al.* (2015, 2016). A 3-km horizontal grid spacing inner domain is nested within a parent grid at 15-km grid spacing that covers the continental United States (Figure 3a) and is run simultaneously in a one-way nested setup. Both domains have 51 identical vertical grid levels, with grid spacing ranging from less than 100 m near the surface to greater than 1 km at the model top. The parent mesoscale ensemble provides the boundary conditions for the nested inner convective-scale ensemble. The placement of the nested domain depends on the severe weather event of interest and covers the region where convection develops. The nested convective-scale domains for the five case studies are shown in Figure 2. A 36-member multi-physics ensemble is initialized at 0000 UTC on the day of the event (Day 1) using analyses from the Global Ensemble Forecast System (Toth *et al.*, 2004; Wei *et al.*, 2008). Different combinations of physics schemes (Table 2) are applied to each ensemble member to account for model physics uncertainties (e.g., Stensrud *et al.*, 2000; Fujita *et al.*, 2007; Yussouf *et al.*, 2015, 2016). Both the outer and inner domains use the same physics options except for the cumulus parametrization scheme, which is turned off in the inner domain. The ensemble system uses the double-moment NSSL microphysics scheme (Mansell *et al.*, 2010).

The GSI system is used for observation preprocessing and calculation of ensemble priors and the EnKF system is used for DA. To create the mesoscale background in the ensemble system, pressure, temperature, water vapor, and horizontal wind components from National Centers for Environmental Prediction (NCEP) "prepared" BUFR (Binary Universal Form for the Representation of meteorological data) format (PrefBUFR) conventional observations are assimilated every hour (Figure 3b) from 0100 UTC on the day of the event (Day 1) to 1000 UTC on the next day (Day 2).

The 1800 UTC convective-scale analyses from the hourly updated system are used as the background (prior) and storm observations are assimilated only in the convective-scale domain using the GSI-EnKF DA system (Johnson *et al.*, 2015; Wang and Wang, 2017). Observations include reflectivity and radial velocity from all operational WSR-88D radars that are located within the nested convective-scale domain (Figure 2), in addition to the aforementioned conventional observations. The hourly updated mesoscale ensemble provides the boundary conditions for the convective-scale system and is cycled every 15 min for a 10-hr period until 0400 UTC the next day (Figure 3c). The radar observation quality control and preprocessing are similar to those of Yussouf *et al.* (2015, 2016). The observation-error standard deviations are assumed to be 5 dBZ for reflectivity and 2 m/s for Doppler velocity. In addition to the relaxation to prior spread (RTPS: Whitaker and Hamill, 2012) technique, which inflates the posterior ensemble spread to a fraction (inflation factor of 0.95 in this study) of the prior ensemble spread at each assimilation cycle, an additive noise technique is used on the ensemble during DA to maintain spread (Dowell and Wicker, 2009).

A continuum of 0–6-hr convective-scale ensemble forecasts are generated from the frequently adjusted ensemble analyses system approximately 1–2 hr prior to the first flash flood report (see Table 1). The evolution of the storms in

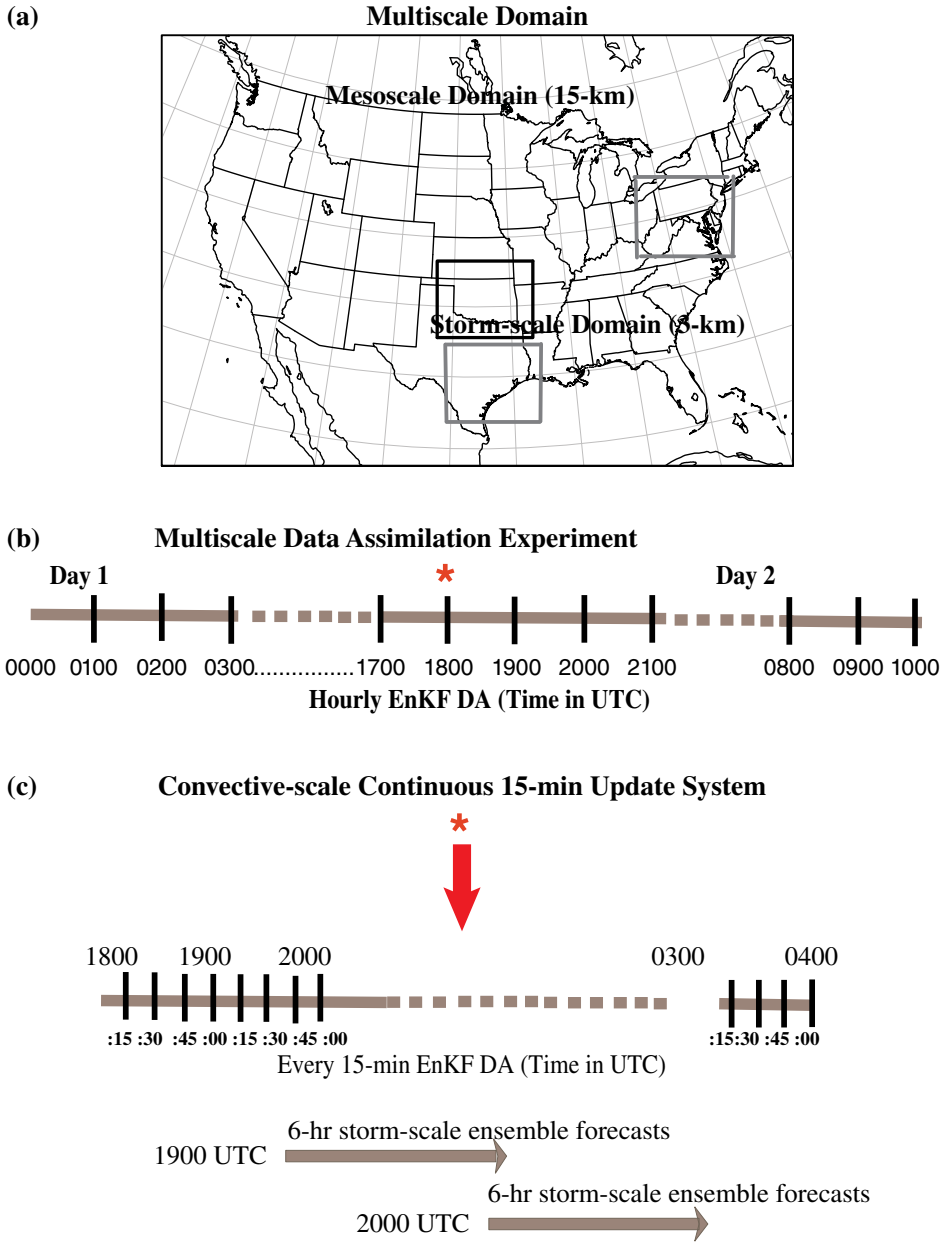

FIGURE 3 (a) The multiscale scale domain with the 15-km horizontal grid-spacing mesoscale domain covering the Continental United States and the nested 3-km convective-scale domain centered over the area of interest. (b) The timeline of the hourly multiscale DA experiments. (c) The timeline for the 15-min cycled convective-scale DA and 0–6-hr ensemble forecast experiments [Colour figure can be viewed at wileyonlinelibrary.com].

the ensemble forecasts is evaluated for the five case studies. Of particular interest in this study is the assessment of how well these 0–6-hr ensemble forecasts simulate heavy rainfall hazards.

4 | RESULTS AND DISCUSSION

4.1 | EnKF performance

The observation-space diagnostics for both reflectivity and Doppler velocity (Yussouf *et al.*, 2013a, 2015, 2016; Wheatley *et al.*, 2015) are calculated to assess the performance of the EnKF system during the 10-hr long continuous radar DA period (Figures 4 and 5). The diagnostics evaluated are the mean innovation (observation – model), total ensemble spread and root-mean squared innovation (*rmsi*), and the consistency ratio from the prior (or background).

The *rmsi*, which is a measure of the overall fit of the observations to the forecasts and analyses, starts with a higher value of 12–14 dBZ for reflectivity, but the error decreases with subsequent assimilation cycles and becomes fairly stable (Figure 4a,c,e,g,i). For Doppler velocity observations, the *rmsi* remains generally the same for the entire DA period

TABLE 2 The physics parametrization options used for the 36-member multiphysics experimental WoFS. Here IC, BC, PBL, SW, and LW stand for initial condition, boundary condition, planetary boundary layer, shortwave, and longwave, respectively

Multiphysics ensemble system							
Ensemble member	**GEFS member for IC and BC**	**Cumulus (15-km grid only)**	**Micro-physics**	**PBL**	**Land surface**	**SW radiation**	**LW radiation**
1	1	Kain–Fritsch	NSSL-2M	YSU	Noah-MP	Dudhia	RRTM
2	2			YSU		RRTMG	RRTMG
3	3			MYJ		Dudhia	RRTM
4	4			MYJ		RRTMG	RRTMG
5	5			MYNN2		Dudhia	RRTM
6	6			MYNN2		RRTMG	RRTMG
7	7	Grell-3	NSSL-2M	YSU	Noah-MP	Dudhia	RRTM
8	8			YSU		RRTMG	RRTMG
9	9			MYJ		Dudhia	RRTM
10	10			MYJ		RRTMG	RRTMG
11	11			MYNN2		Dudhia	RRTM
12	12			MYNN2		RRTMG	RRTMG
13	13	Tiedke	NSSL-2M	YSU	Noah-MP	Dudhia	RRTM
14	14			YSU		RRTMG	RRTMG
15	15			MYJ		Dudhia	RRTM
16	16			MYJ		RRTMG	RRTMG
17	17			MYNN2		Dudhia	RRTM
18	18			MYNN2		RRTMG	RRTMG
19	18	Kain–Fritsch	NSSL-2M	YSU	Noah-MP	Dudhia	RRTM
20	17			YSU		RRTMG	RRTMG
21	16			MYJ		Dudhia	RRTM
22	15			MYJ		RRTMG	RRTMG
23	14			MYNN2		Dudhia	RRTM
24	13			MYNN2		RRTMG	RRTMG
25	12	Grell	NSSL-2M	YSU	Noah-MP	Dudhia	RRTM
26	11			YSU		RRTMG	RRTMG
27	10			MYJ		Dudhia	RRTM
28	9			MYJ		RRTMG	RRTMG
29	8			MYNN2		Dudhia	RRTM
30	7			MYNN2		RRTMG	RRTMG
31	6	Tiedke	NSSL-2M	YSU	Noah-MP	Dudhia	RRTM
32	5			YSU		RRTMG	RRTMG
33	4			MYJ		Dudhia	RRTM
34	3			MYJ		RRTMG	RRTMG
35	2			MYNN2		Dudhia	RRTM
36	1			MYNN2		RRTMG	RRTMG

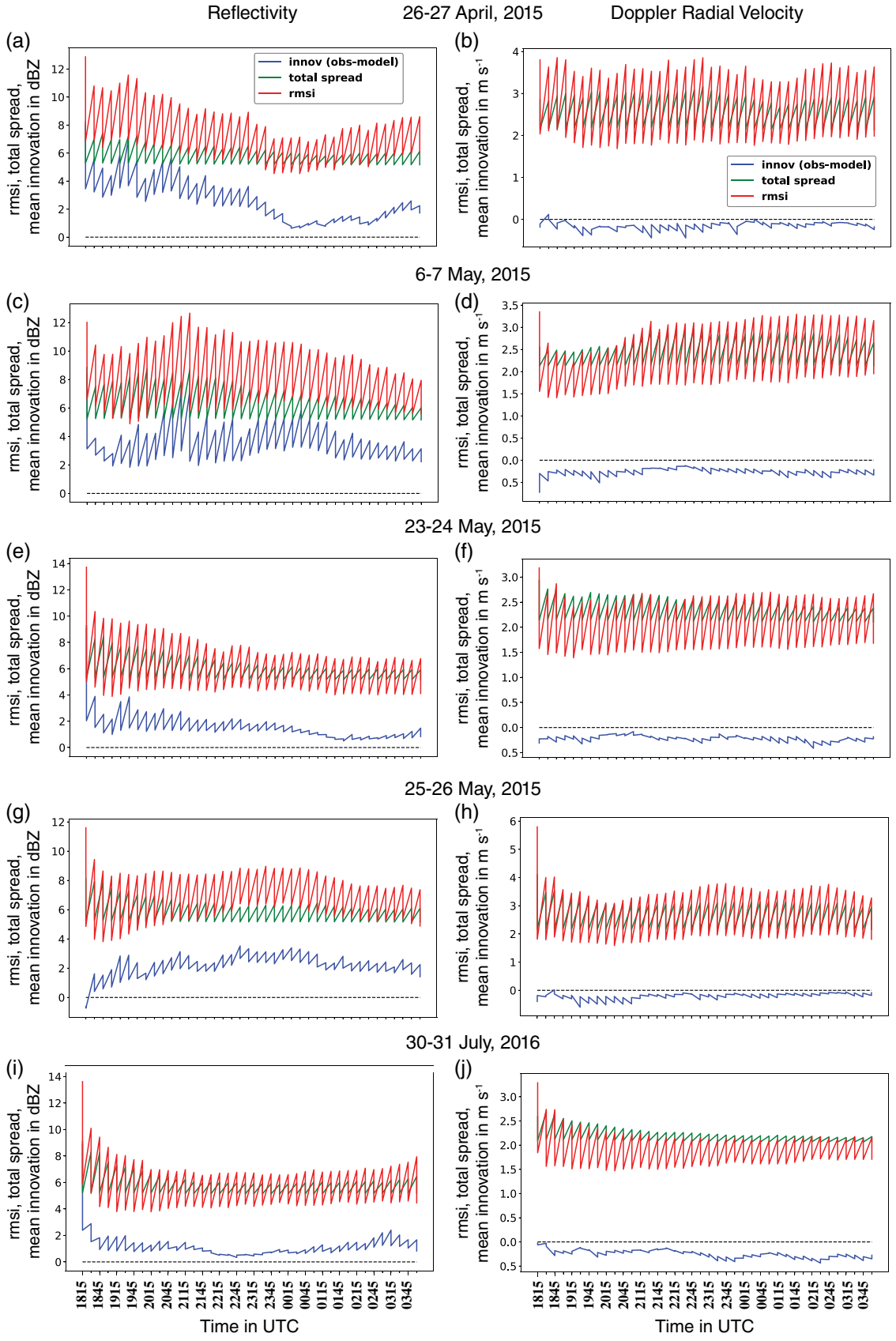

FIGURE 4 The observation-space diagnostic statistics of root-mean-square innovation (rmsi), total ensemble spread and mean innovations for assimilated (a, c, e, g, i) reflectivity (dBZ) and (b, d, f, h, j) Doppler velocity (m/s) observations, respectively, during the 10-hr DA period in the convective-scale domain for the five case studies. The sawtooth patterns are due to the plotted forecast and analysis statistics [Colour figure can be viewed at wileyonlinelibrary.com].

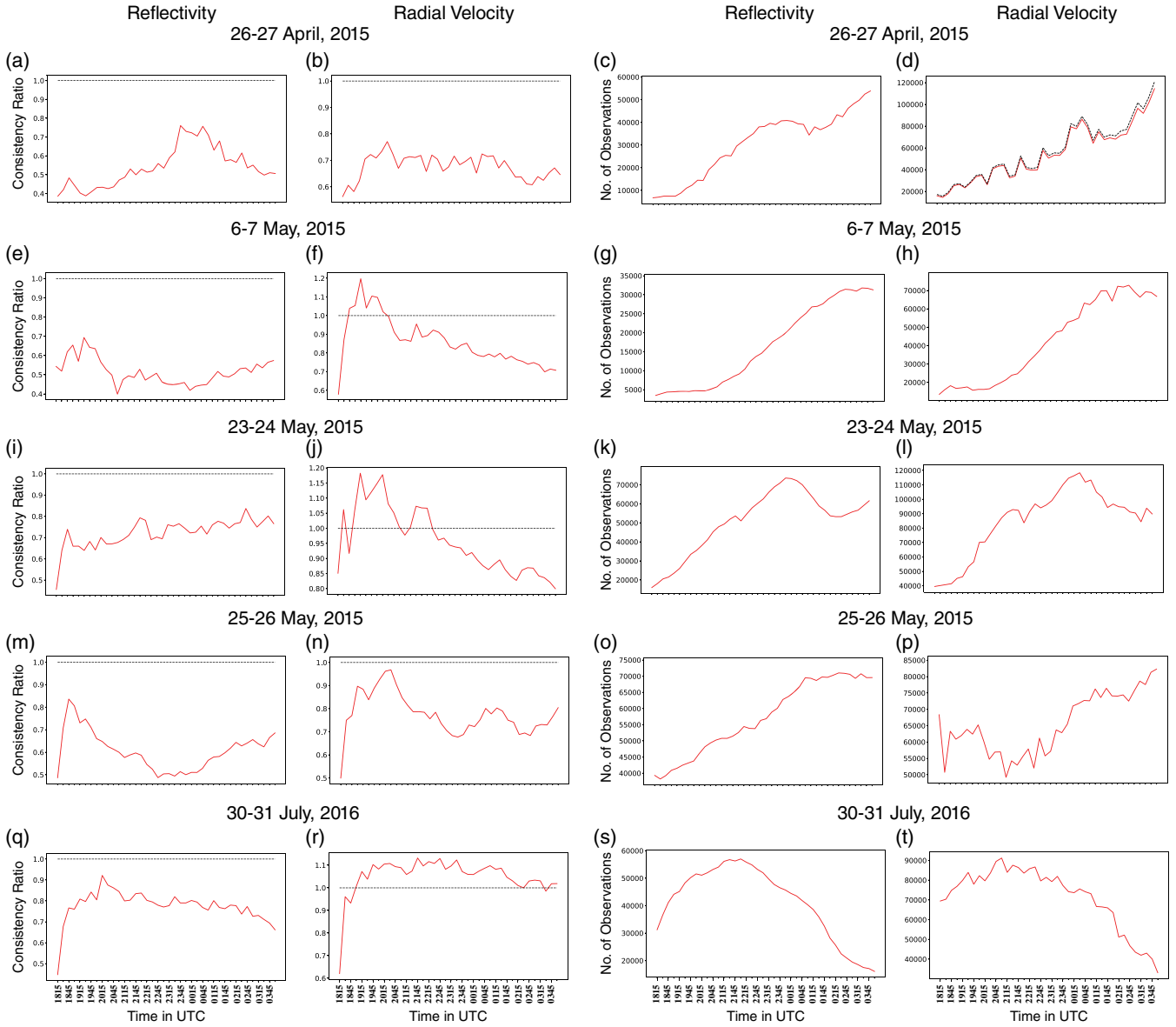

FIGURE 5 The consistency ratio calculated from the (left two columns) prior (forecast background) and the (right two columns) total number of reflectivity and Doppler velocity observations assimilated during the 10-hr every 15 min radar DA period in the convective-scale domain for the five case studies [Colour figure can be viewed at wileyonlinelibrary.com].

(Figure 4b,d,f,h,j). The total spread and the *rmsi* are reasonably similar in magnitude for both Doppler velocity and reflectivity, indicating that the ensemble spread is representative of the forecast error for the 15-min cycle of the WoFS.

The mean innovation measures the forecasts and analysis bias. The mean innovation is positive for reflectivity, which reveals that the ensemble system underpredicts reflectivity during the entire assimilation period (Figure 4). The mean innovation for reflectivity starts with a higher value, which is largely due to initial spin-up of storms in the model but decreases with time, and varies within the range 0.5–2.0 dBZ during the later time of the DA period. For Doppler velocity, the mean innovation is very close to 0 (-0.5 ms^{-1} or less) during the entire DA window for all five cases. The slight overprediction, also seen in previous studies (Yussouf *et al.*, 2015, 2016), is likely attributable to differences in observation geometry and model grid spacing, model physics, precipitation fall speed or a combination of all.

For the assumed observation error, if the prior ensemble variance is a good approximation of the forecast-error variance, the consistency ratio should be close to 1.0. The consistency ratio for reflectivity remains within the range 0.4–0.9 (Figure 5a,e,i,m,q). The initial value of the consistency ratio for the Doppler velocity is 0.5, which increases rapidly with the continuous DA cycles to magnitudes in the range of 0.8–1.2 (Figure 5b,f,j,n,r). Although the values of the consistency ratio are in a good range, the total spread in each case remains rather constant in time, particularly for reflectivity. Future work will incorporate and test other variants of

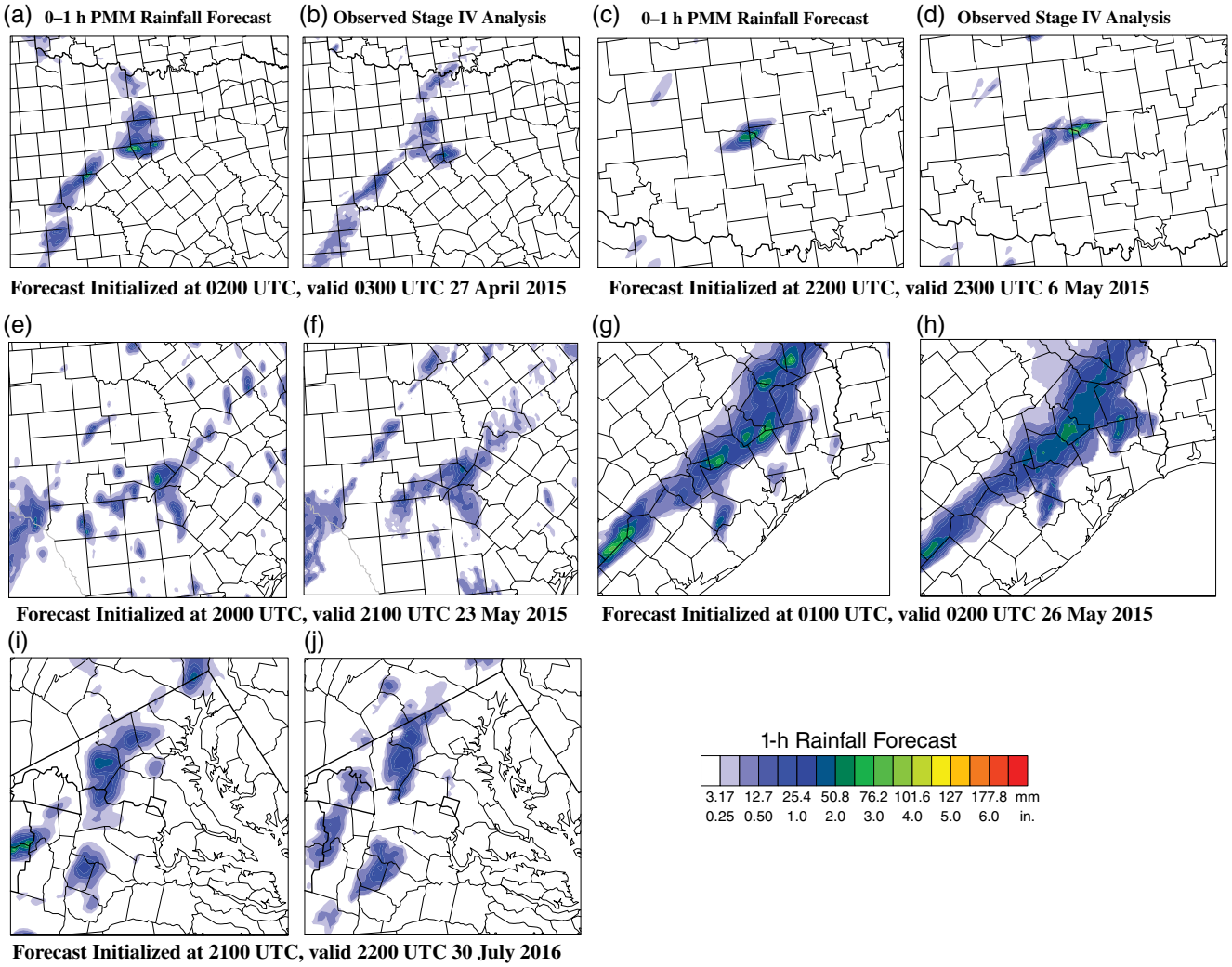

FIGURE 6 (a, c, e, g, i) Probability-matched mean (PMM) 0–1-hr rainfall forecasts (in mm and inches, see color bar) for the five cases, as well as (b, d, f, h and j) NCEP's Stage-IV 1-hr accumulated rainfall analyses valid at the same time. A neighborhood radius of 15 km is used to calculate the mean. The portion of the domain shown covers the areas of interest [Colour figure can be viewed at wileyonlinelibrary.com].

the inflation technique for the convective scale: for example, spatially and temporally varying adaptive inflation (Anderson, 2009) in the GSI-EnKF DA system, for a more realistic ensemble spread during the life cycle of the storms. The number of reflectivity (Figure 5c,g,k,o,s) and radial velocity (Figure 5d,h,l,p,t) observations assimilated varies from case to case and depends on the maturity of the storms. The overall quality of the observation-space diagnostics (Figures 4 and 5) for reflectivity and radial velocity demonstrates that the DA is tuned reasonably well for a fairly stable system. Importantly, there is no indication of filter divergence during the 10-hr long 15-min cycled assimilation period.

4.2 | Probability-matched mean rainfall forecast

Unlike the traditional ensemble mean, the probability-matched mean (PMM: Ebert, 2001) helps to retain the high amplitude characteristics of the individual ensemble members. The regridded[1] WoFS members are used to calculate the PMM using a neighborhood approach (Ebert, 2009) with a radius of 15 grid points for 0–1-hr, 0–3-hr, and 0–6-hr accumulation periods (Figures 6–8). The PMM rainfall forecasts are compared with the NCEP Stage IV analysis (Baldwin and Mitchell, 1997; Lin and Mitchell, 2005). The Stage IV analysis consists of mosaicked national 1-hr and 6-hr precipitation estimates from the multi-sensors (radar and rain-gauge measurements) produced by the 12 River Forecast Centers (RFCs). The rainfall forecasts shown in the plots are initialized such that the forecast covers the time period of the observed intense rainfall from the events. Visual inspection shows that the 0–1 hr PMM rainfall forecasts (Figure 6) are remarkably similar to the observed stage IV rainfall for all five cases. The explicit initialization of the storms in

[1]The 3-km horizontal grid spacing ensemble members are regridded to the ~4.7-km grid spacing Stage-IV analysis using the neighbor-budget interpolation technique (e.g., Accadia *et al.,* 2003).

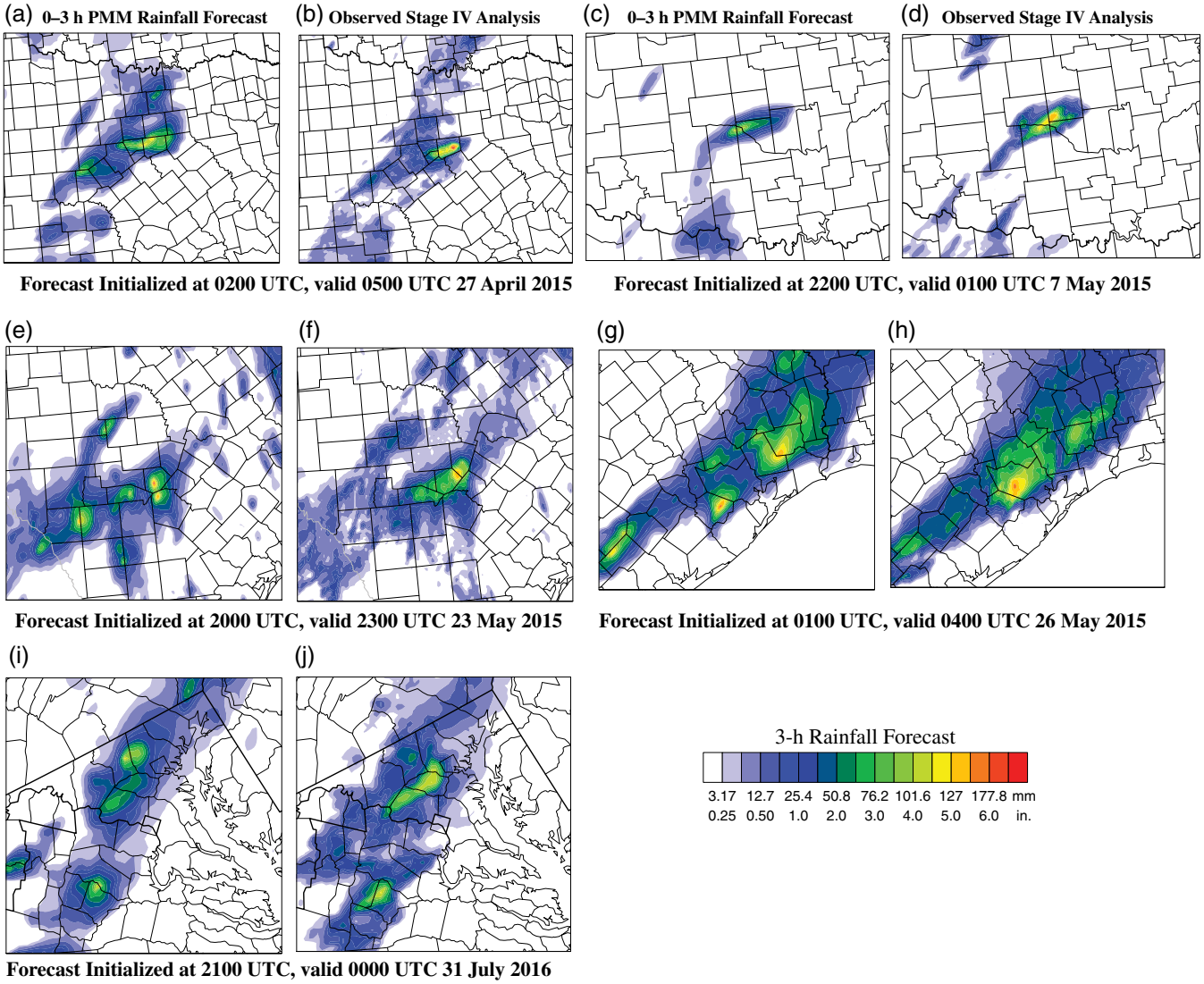

FIGURE 7 Same as in Figure 6, but for 0–3-hr accumulated rainfall forecasts [Colour figure can be viewed at wileyonlinelibrary.com].

the ensemble system, owing to the frequent assimilation of radar observations, clearly reduced the uncertainty of the rainfall. While the WoFS seems to capture the most intense rainfall reasonably well, there are differences in the location, areal coverage, and magnitude of accumulated rainfall as the forecast lead time increases (Figures 7 and 8). For example, compared with the observed precipitation (Figure 7b,d), the heaviest rainfall amounts are slightly underpredicted in the 3-hr PMM forecast (Figure 7a,c). The 6-hr PMM rainfall for the May 23–24, 2017 Texas event (Figure 8e,f) predicts an additional small area of high precipitation just south of the main rainfall core, where stage IV shows much lighter rainfall.

4.3 | Ensemble-derived exceedance probabilities of rainfall forecasts

The WoFS exceedance probabilities of rainfall totals are calculated at raw model grid points and rainfall thresholds ranging from 12.7 mm (0.5 inches) to 76.2 mm (3.0 inches) are chosen to highlight the areas of most intense rainfall (Figures 9–11), which can lead to flash flood. The thick black contours overlaid are the observed Stage-IV rainfall for the threshold amounts indicating the area. The black dots are the NWS flash flood reports from *storm data* valid during the forecast period. The location and timing of the flash flood reports give an idea of the areal coverage and timing where heavy rainfall led to flash flooding. Even though the flash flood reports from the storm data are useful for validating ensemble-derived forecast products, subjectivity due to the human element in the flash-flood reporting process must be taken into consideration when interpreting the results.

The 1- and 3-hr rainfall forecasts generate 100% probabilities for the dominant precipitation core in the observed rainfall area with good accuracy (Figures 9 and 10) and the probability values are located closer to the flash flood reports. There are a few areas where the WoFS indicates no rainfall forecast in the observed location, for example to the south of the high rainfall core for the April 26–27, 2015 case (Figure 10a).

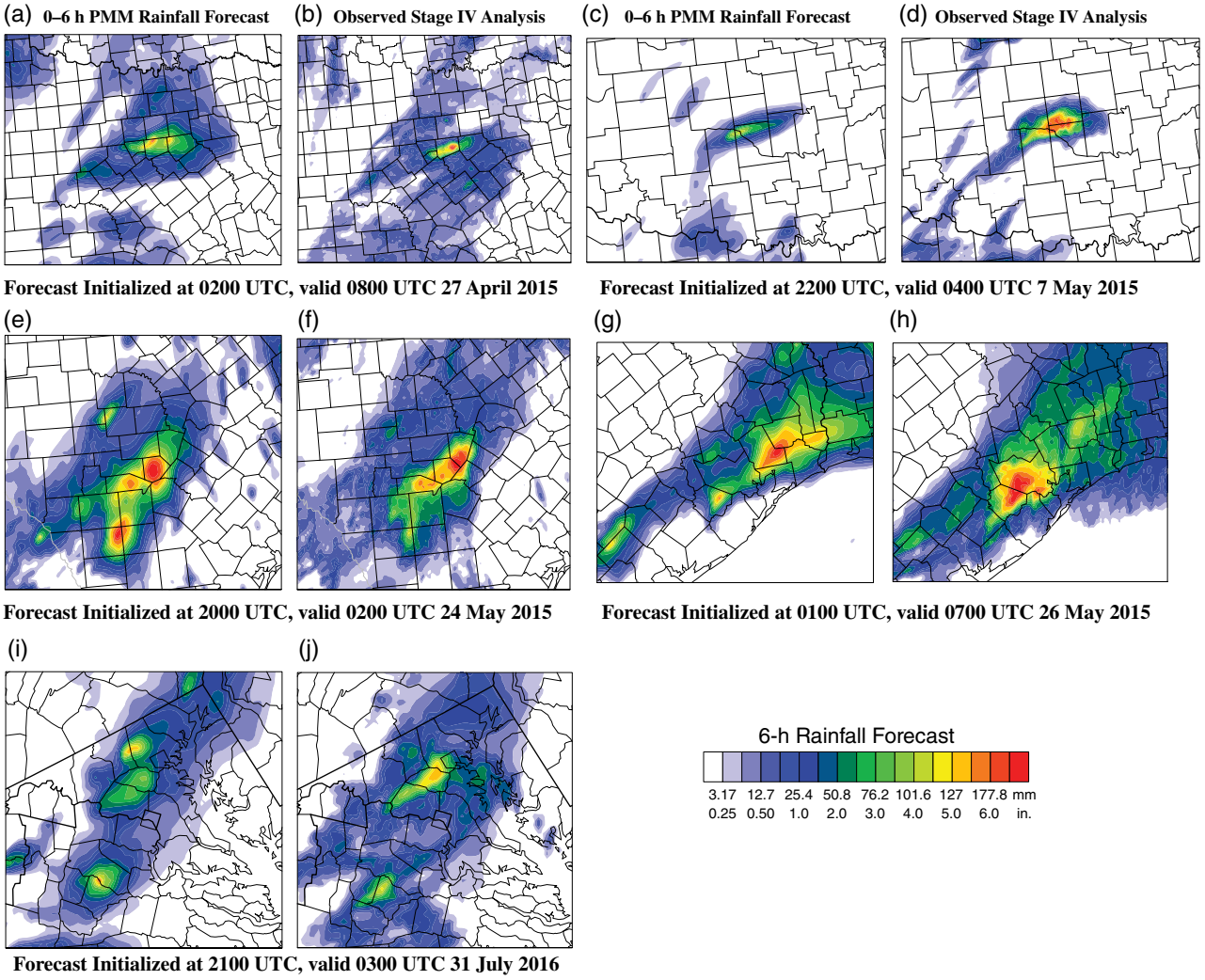

FIGURE 8 Same as in Figure 6, but for 0–6-hr accumulated rainfall forecasts [Colour figure can be viewed at wileyonlinelibrary.com].

During the longer 6-hr accumulation period (Figure 11), the WoFS forecasts a broad swath of higher probability values, in particular for the 50.8 mm threshold (Figure 11a,c,e,g,i). Most of the storm data flash flood reports are within the ensemble probability envelope. However, closer inspection reveals that the predicted 100% probabilities are at times smaller than the observed rainfall cores, indicating ensemble underdispersion later in the forecast period. Better design of convective-scale ensembles for 0–6-hr forecasts will likely alleviate the underdispersion of the system. Some rainfall cores are even associated with zero probabilities. Factors likely to play a role include deficiencies in parametrization schemes for convective-scale modeling. Furthermore, there are some displacement errors to the eastnortheast compared with the stage IV rainfall contours, in particular for the May 6–7, 2015 central Oklahoma and May 26–27, 2015 southwest Texas events (Figure 11c,d,g,h). The displacement is due to the tendency of the forecast convective system to move faster than the actual system, which is a common artifact of convective-scale modeling (Snook *et al.*, 2015; Wheatley *et al.*, 2015; Yussouf *et al.*, 2015).

Additional insight into the ability of WoFS to predict flash flood potential is assessed by calculating the ensemble-derived probability of 0–6-hr rainfall forecasts exceeding NWS flash flood guidance (FFG: Schmidt *et al.*, 2007; Clark *et al.*, 2014). The FFG values are estimated by the 12 RFCs over the conterminous United States to aid NWS forecasters in monitoring and providing warning of impending flash flood threats. The FFG products, which are the estimate of rainfall necessary over the next 6 hr to cause rivers and small streams to overflow their natural banks, are issued at 0000, 0600, 1200, and 1800 UTC each day. The FFG products take soil moisture and stream-flow conditions into consideration. The WoFS initialized at 0000 UTC is used to calculate 0–6-hr accumulated rainfall probabilities exceeding FFG issued at 0000 UTC. The WoFS rainfall probability values exceeding critical rainfall thresholds to cause bank-full conditions (i.e., FFG) for all five case studies are shown in Figure 12. The high probability values for the April 26–27, 2015 north–central Texas, May 6–7, 2015 central Oklahoma, May 26–27, 2015 southwest Texas, and July 30–31, 2016 Maryland events correspond closely with the flash flood

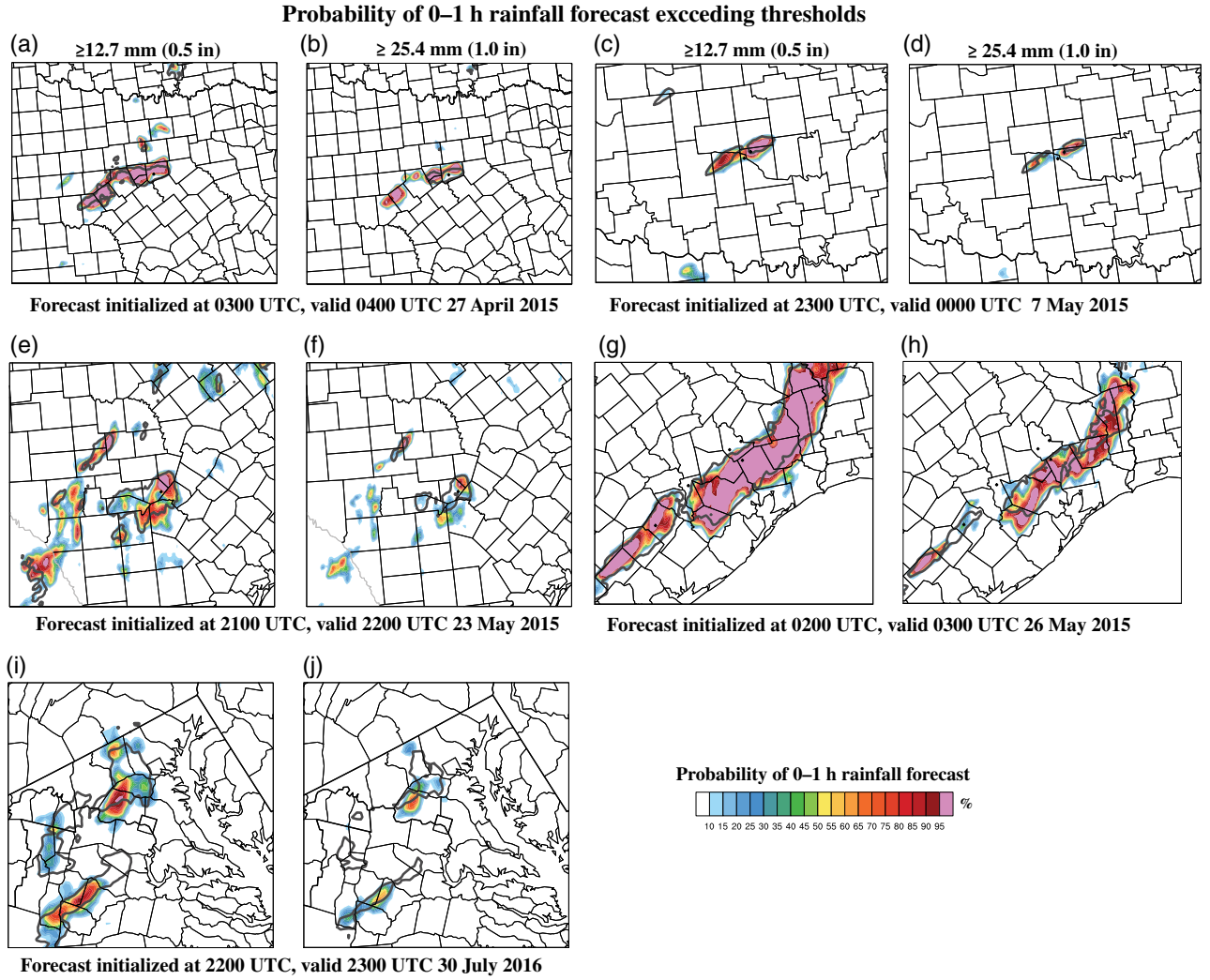

FIGURE 9 Ensemble-derived probability of 0–1-hr rainfall forecasts greater than (a, c, e, g, i) 12.7 mm (0.5 inches; colors, 5% increment) and (b, d, f, h, j) 25.4 mm (1.0 inch) for the five cases. The thick black contour overlaid is the Stage-IV rainfall for the threshold values. The black dots are the NWS flash flood reports from storm data during the forecast time period. The portion of the domain shown covers the areas of interest [Colour figure can be viewed at wileyonlinelibrary.com].

reports (Figure 12a,b,d,e). The WoFS forecast probabilities are comparatively low for the May 23–24, 2015 south–central Texas event, which do receive as much heavy rainfall but are still flooded according to the flash flood reports (Figure 12c). While the probability of rainfall forecasts exceeding the FFG can provide insight into the ability of the WoFS in flash flood prediction, caution should be taken when comparing the FFG with WoFS rainfall amounts, since they are not the same and thus may not always match well. Nevertheless, the overall forecast probabilities of exceeding FFG clearly show the promise of the WoFS for short-term flash flood threats.

4.4 | Fractions skill scores

The fractions skill scores (FSSs: Duc *et al.*, 2013; Schwartz *et al.*, 2010; Roberts and Lean, 2008) for hourly aggregated rainfall from the five cases are calculated for quantitative verification of the probabilistic forecasts. The FSS measures the spatial skill of rainfall, with a score of 1 indicating perfect skill. The method used to calculate FSSs from the ensemble members follows Schwartz *et al.* (2010; using their equations 6–8), which is an extension of the FSS in Roberts and Lean (2008) for ensemble-based probabilistic forecast verification. The thresholds for hourly rainfall during the 0–6-hr forecast period range from 12.7 mm (0.5 inches) to 76.2 mm (3 inches). The FSSs are calculated using a neighborhood approach (Ebert, 2009), with several neighborhood radii ranging from 3–36 km (Figure 13) over the same domain and time period as in Figures 6–11, using hourly Stage IV rainfall as the observations.

The FSSs are considerably larger for smaller rainfall thresholds (i.e., 12.7 and 25.4 mm, Figure 13a,b) than for higher thresholds (i.e., 50.8 and 76.2 mm, Figure 13c,d), indicating that the system is less skillful in forecasting the location of the most intense rainfalls. Higher rainfall amounts

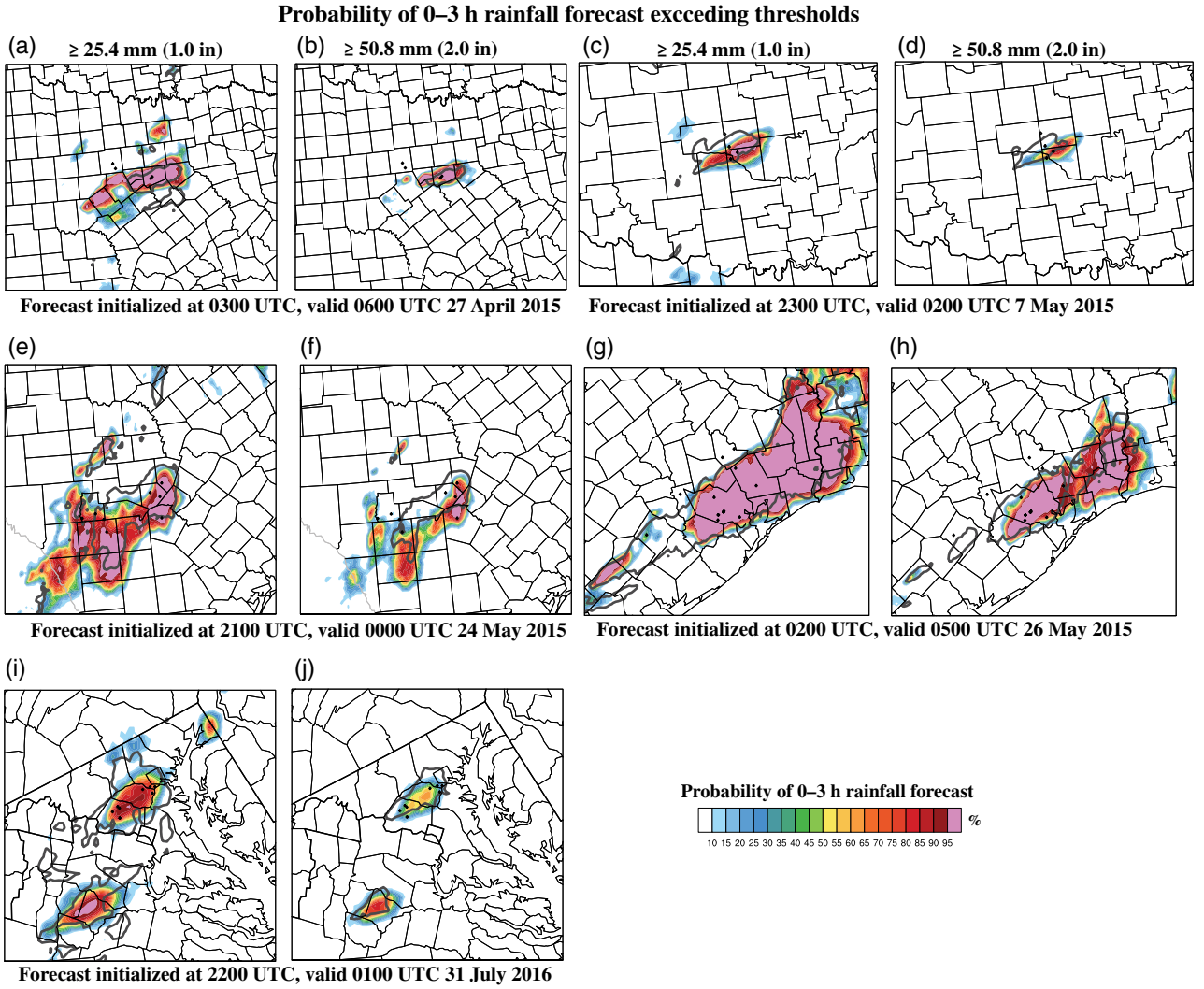

FIGURE 10 Same as in Figure 9, but for ensemble probability of 0–3-hr rainfall forecasts exceeding thresholds 25.4 mm (1.0 inch) and 50.8 mm (2.0 inches) [Colour figure can be viewed at wileyonlinelibrary.com].

highlight smaller precipitation cores in both observations and in forecasts, which likely reduces the overlap between the two due to position and magnitude errors (Baldwin and Kain, 2006; Yussouf *et al.*, 2016). Consistent with the subjective assessment from the probability forecast, the skill scores decrease at longer forecast lead times, in particular after 3–4 hr. Not surprisingly, the FSSs increase as the neighborhood radii increase from 3–36 km. As the neighborhood increases, more ensemble members overlap or agree, which leads to higher FSS scores.

4.5 | Ensemble maximum rainfall from July 30, 2016 Ellicott City flash flood

The historic flash flood that occurred in Ellicott City, Maryland was highly localized in nature. The ELYM2 rain gauge, which is part of the Hydrometeorological Automated Data System (HADS), located in Ellicott City measured more than 160 mm (6.30 inches) of rainfall over a 3-hr period (Figure 14a). To estimate the WoFS predicted rainfall at the ELYM2 rain-gauge location, the maximum rainfall during the same 3-hr period (forecasts are initialized at 2200 UTC) is calculated from the ensemble members within a 9-km neighborhood radius from the rain-gauge location. The WoFS rainfall forecasts (Figure 14b) are of comparable magnitude with the rain-gauge measured amounts (Figure 14a) during the 3-hr accumulation period. The rain gauge measured a total rainfall of 160.53 mm (6.32 inches) at the end of the 3-hr accumulation period, whereas the WoFS forecast accumulated rainfall at the end of same period is 148.84 mm (5.86 inches).

5 | SUMMARY AND CONCLUDING REMARKS

The NOAA's WoF project is developing an on-demand, high-resolution, ensemble-based, 0–6-hr NWP modeling system capable of supporting NOAA's watch-to-warning

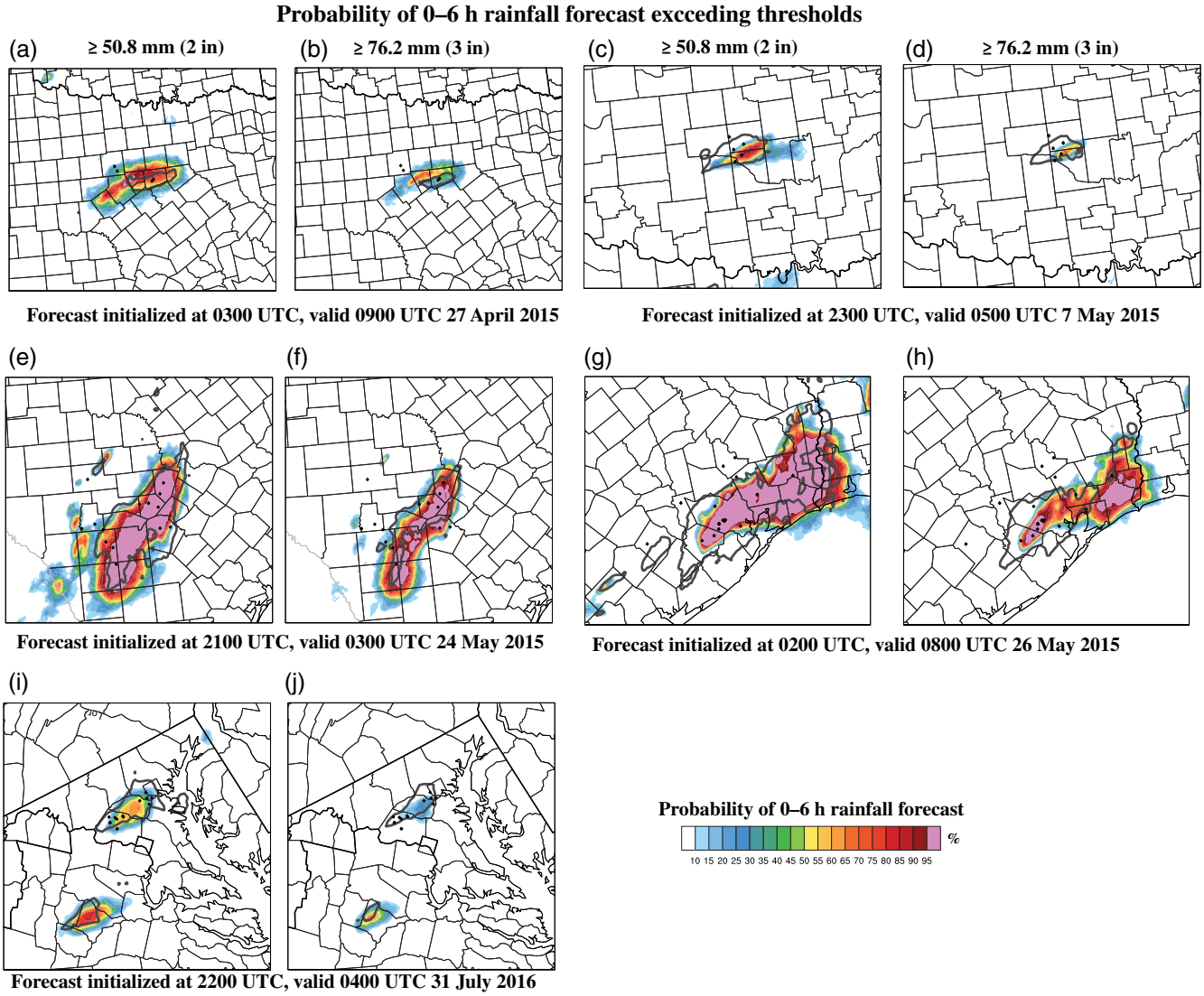

FIGURE 11 Same as in Figure 9, but for ensemble probability of 0–6-hr rainfall forecasts exceeding thresholds 50.8 mm (2.0 inches) and 76.2 mm (3.0 inches) [Colour figure can be viewed at wileyonlinelibrary.com].

operations of severe convective storms and associated threats. The WoFS is a frequently cycled, convective-scale, ensemble DA and prediction system. Multiple studies have shown the potential of the WoFS for 0–1-hr probabilistic forecasts of severe thunderstorms and low-level rotations (Yussouf et al., 2013a, 2013b; 2015; Wheatley *et al.*, 2015; Jones *et al.*, 2016; Skinner *et al.*, 2016). The system also demonstrated skillful rainfall forecasts of an intense convective rainfall and flash flood event (Yussouf *et al.*, 2016). Accurate short-term heavy rainfall forecasts are crucial for improving flash flood warnings and precipitation nowcasting. To assess the system's robustness to produce skillful 0–6-hr probabilistic forecasts of heavy rainfall events that often lead to flash flooding, multiple significant events from the 2015 and 2016 warm seasons are simulated in this study. Several WSR-88Ds reflectivity and radial velocity observations and other routinely available conventional observations are assimilated every 15 min into a 36-member multiphysics, experimental WoFS at 3-km horizontal grid spacing. The system uses the WRF-ARW model and NWS operational GSI-EnKF DA system. A series of 0–6-hr ensemble forecasts is initialized from the frequently adjusted ensemble analyses at the top of each hour, starting around 1–2 hr prior to the first flash flood report.

The GSI-EnKF diagnostic statistics indicate stability during the 10-hr long cycled assimilation window, indicating the DA system is tuned reasonably well for radar reflectivity and radial velocity observations. The PMM rainfall from all five experiments is able to reproduce the convection and the area of most intense rainfall with reasonable accuracy. In particular, the 0–1-hr PMM rainfall forecasts matches NCEP's Stage-IV analyses very well, suggesting that the convection is initialized well in the ensemble. The ensemble-derived probabilistic rainfall forecasts are skillful during 0–3 hr for the intense rainfall that overlaps Stage-IV analyses and storm data flash flood reports with reasonable accuracy in location and

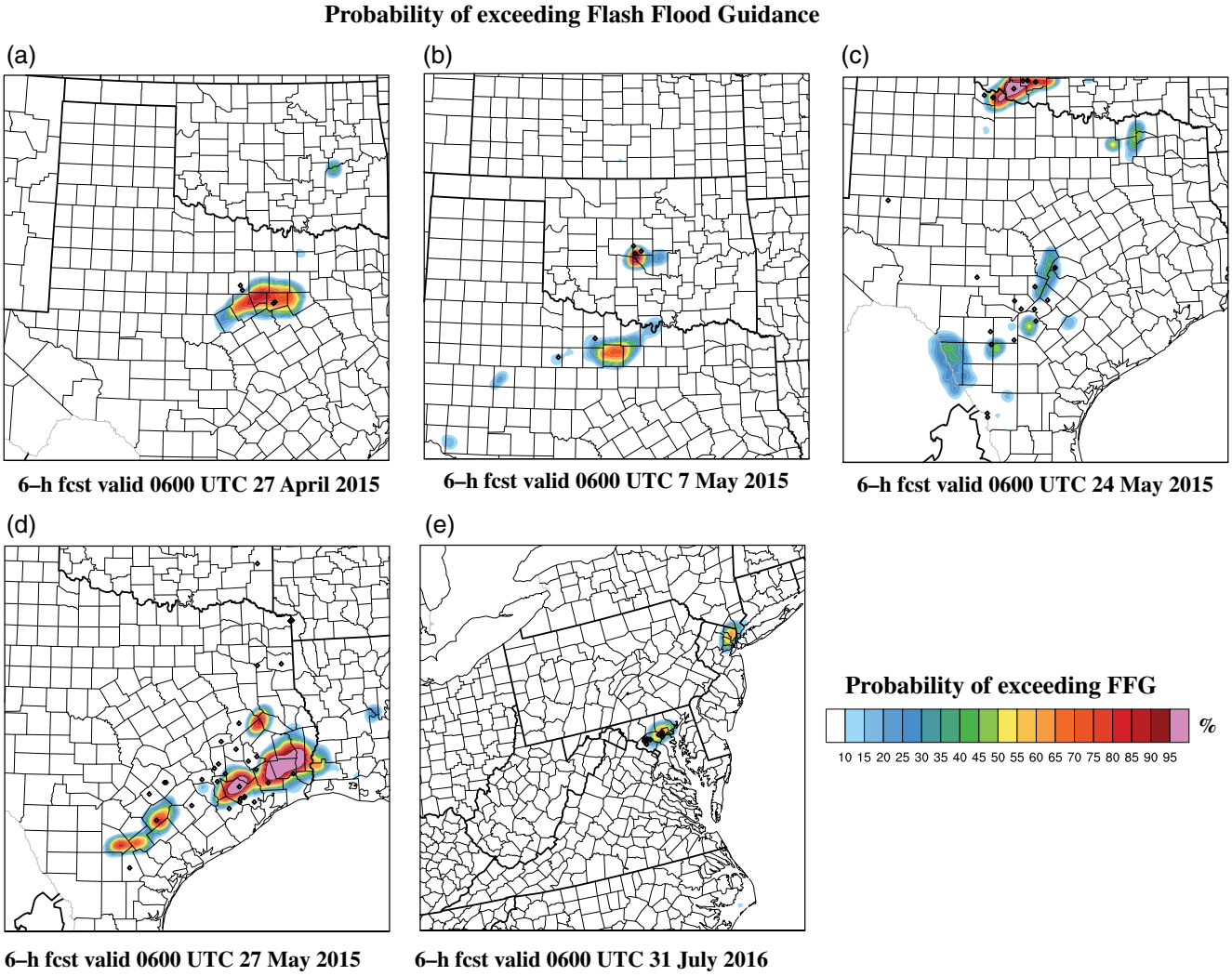

FIGURE 12 Ensemble-derived exceedance probability of 0–6-hr rainfall forecasts greater than the flash flood guidance (colors, 5% increment). The black dots are the NWS flash flood reports from storm data during the forecast time period [Colour figure can be viewed at wileyonlinelibrary.com].

amount of rainfall. The ensemble-derived forecast probability of rainfall where FFG is being met or exceeded also signifies the potential of the system to highlight areas of flash flood threats. However, compared with the observed precipitation, the ensemble forecast for the longer 6-hr accumulation period shows noticeable northeastward displacement in the heavy precipitation core, with larger areal coverage of 100% probabilities, and misses several observed heavy rainfall cores where the forecast probability is 0%. These findings demonstrate the challenges associated with developing a WoF type system and motivate us to conduct future studies to improve the ensemble configuration to better represent typical 0–6-hr forecast errors and reduce model errors. One possible approach to add sufficient spread in WoFS is to incorporate a stochastically perturbed physics parametrization in the ensemble design (Jankov *et al.*, 2019 and references therein). The quantitative FSS verification metric also paints a consistent picture with the subjective assessments of the probability forecasts. Not surprisingly, the FSSs are higher during the ~0–3-hr forecast period and the skill decreases during the later forecast period.

Overall results show the potential of the WoFS in forecasting short-term heavy rainfall with higher accuracy and specificity. The continuum of these frequently updated 0–6-hr rainfall forecasts can increase situational awareness and warning lead times of heavy rainfall threats. The current state-of-the-art distributed hydrologic models use observed rainfall rates as the forcing mechanism to predict stream discharge products (Gourley *et al.*, 2017) and thus do not provide sufficient warning lead times for flash flood threats. Accurate 0–6-hr rainfall products from a high-resolution WoFS can provide a pathway for extending explicit flash flood prediction lead times by using a forcing mechanism on the distributed hydrologic model. Future work will evaluate the potential of an integrated WoF–hydrologic modeling system to increase probabilistic flash flood warning lead time.

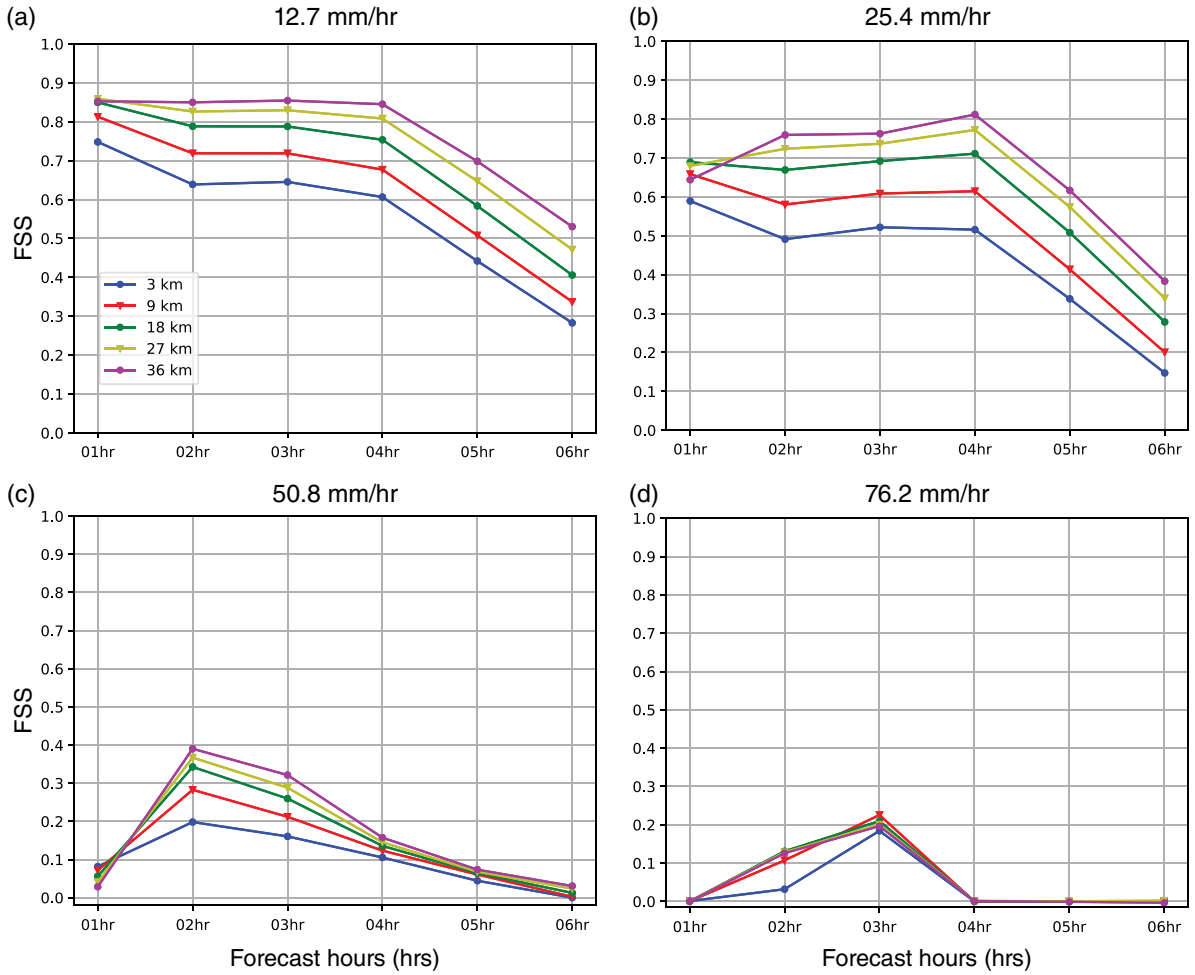

FIGURE 13 The fraction skill score (FSS) as a function of forecast hour for (a) 12.7 mm (0.5 inches), (b) 25.4 mm (1.0 inch), (c) 50.8 mm (2.0 inches), and (d) 76.2 mm (3.0 inches) accumulated rainfall thresholds. The aggregated FSSs from the five cases are calculated from the ensemble forecast sets and the domain sizes as in Figures 6–11, using different neighborhood thresholds. Details are shown in the legend [Colour figure can be viewed at wileyonlinelibrary.com].

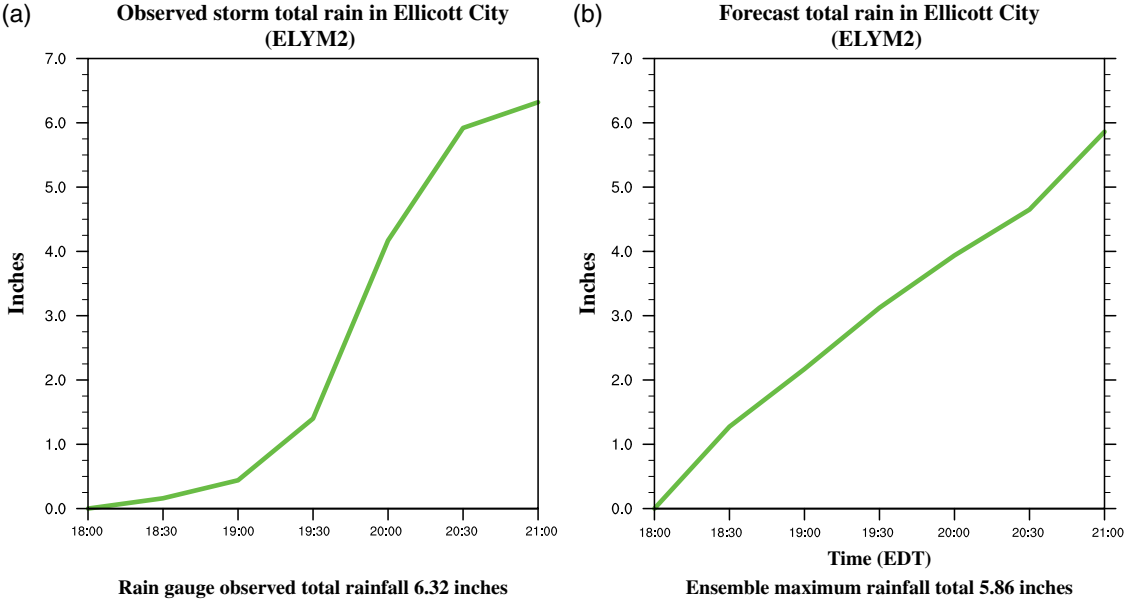

FIGURE 14 (a) The observed total rainfall measured by the ELYM2 rain gauge in Ellicott City, Maryland, and (b) the forecast ensemble maximum rainfall total at the same location from 2200 UTC on July 30, 2016 to 0100 UTC on July 31, 2016 (1800–2100 Eastern Daylight Time). The WoFS is initialized at 2200 UTC [Colour figure can be viewed at wileyonlinelibrary.com].

ACKNOWLEDGEMENTS

We thank the two anonymous reviewers for providing insightful comments that led to substantial improvement of the manuscript. We also thank Anthony Reinhart for helping with processing the radar data and Patrick Skinner for providing the PMM code and helpful discussions. Stage-IV data are provided by NCAR/EOL under the sponsorship of the National Science Foundation (https://data.eol.ucar.edu/). The supercomputer where all the experiments are run is maintained by Gerry Creager, Brett Morrow, Steven Fletcher, and Robert Coggins. Funding for this research was provided by the NOAA/NSSL FY 2016 Director's Discretionary Research Fund and NOAA–University of Oklahoma Cooperative Agreement #NA11OAR4320072, U.S. Department of Commerce.

ORCID

Nusrat Yussouf https://orcid.org/0000-0003-4998-1770

REFERENCES

Accadia, C., Mariani, S., Casaioli, M., Lavagnini, A. and Speranza, A. (2003) Sensitivity of precipitation forecast skill scores to bilinear interpolation and a simple nearest-neighbor average method on high-resolution verification grids. *Weather and Forecasting*, 18, 918–932.

Anderson, J.L. (2009) Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus A*, 61, 72–83.

Baldwin, M.E. and Kain, J.S. (2006) Sensitivity of several performance measures to displacement error, bias, and event frequency. *Weather and Forecasting*, 21, 636–648. https://doi.org/10.1175/WAF933.1.

Baldwin, M.E. and Mitchell, K.E. (1997) The NCEP hourly multisensory U.S. precipitation analysis for operations and GCIP research. *Preprints, 13th Conference on Hydrology*. Long Beach, CA: American Meteorological Society, pp. 54–55.

Clark, R.A., Gourley, J.J., Flamig, Z.L., Hong, Y. and Clark, E. (2014) CONUS-wide evaluation of National Weather Service flash flood guidance products. *Weather and Forecasting*, 29, 377–392.

Dawson, D.T., Wicker, L.J., Mansell, E.R. and Tanamachi, R.L. (2012) Impact of the environmental low-level wind profile on ensemble forecasts of the 4 May 2007 Greensburg, KS tornadic storm and associated mesocyclones. *Monthly Weather Review*, 140, 696–716.

Developmental Testbed Center (2017a) *Gridpoint Statistical Interpolation User's Guide Version 3.6,* 158 pp. Available at: https://dtcenter.org/com-GSI/users/docs/ [Accessed 15th September 2018].

Developmental Testbed Center (2017b) *Ensemble Kalman Filter (EnKF) User's Guide for Version 1.2*, 86 pp. Available at: http://www.dtcenter.org/EnKF/users/docs/index.php [Accessed 15th September 2018].

Dowell, D.C. and Wicker, L.J. (2009) Additive noise for storm-scale ensemble forecasting and data assimilation. *Journal of Atmospheric and Oceanic Technology*, 26, 911–927.

Duc, L., Saito, K. and Seko, H. (2013) Spatial–temporal fractions verification for high resolution ensemble forecasts. *Tellus*, 65A, 18171.

Ebert, E.E. (2001) Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Monthly Weather Review*, 129, 2461–2480. https://doi.org/10.1175/1520-3380493(2001)1292.0.CO;2.

Ebert, E.E. (2009) Neighborhood verification: a strategy for rewarding close forecasts. *Weather and Forecasting*, 24, 1498–1510.

Fujita, T., Stensrud, D.J. and Dowell, D.C. (2007) Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physics uncertainty. *Monthly Weather Review*, 135, 1846–1868. https://doi.org/10.1175/MWR3391.1.

Gourley, J.J., Flamig, Z.L., Vergara, H., Kirstetter, P.-E., Clark, R.A., Argyle, E., Arthur, A., Martinaitis, S., Terti, G., Erlingis, J.M., Hong, Y. and Howard, K. (2017) The FLASH project: improving the tools for flash flood monitoring and prediction across the United States. *Bulletin of the American Meteorological Society*, 98, 361–372.

Houtekamer, P.L., Mitchell, H.L., Pellerin, G., Buehner, M., Charron, M., Spacek, L. and Hansen, B. (2005) Atmospheric data assimilation with an Ensemble Kalman Filter results with real observations. *Monthly Weather Review*, 133, 604–620.

Jankov, I., Beck, J., Wolff, J.K., Harrold, M., Olson, J.B., Smirnova, T., Alexander, C. and Berner, J. (2019) Stochastically perturbed parametrizations in an HRRR-based ensemble. *Monthly Weather Review*, 147, 153–173.

Johnson, A., Wang, X., Carley, J.R., Wicker, L.J. and Karstens, C. (2015) A comparison of multiscale GSI-based EnKF and 3DVar data assimilation using radar and conventional observations for midlatitude convective-scale precipitation forecasts. *Monthly Weather Review*, 143, 3087–3108.

Jones, T.A., Knopfmeier, K., Wheatley, D.M., Creager, G., Minnis, P. and Palikondo, R. (2016) Storm-scale data assimilation and ensemble forecasting with the NSSL experimental Warn-on-Forecast system. Part II: combined radar and satellite data experiments. *Weather and Forecasting*, 31, 297–327. https://doi.org/10.1175/WAF-D-15-0107.1.

Kleist, D.T., Parrish, D.F., Derber, J.C., Treadon, R., Wu, W.-S. and Lord, S. (2009) Introduction of the GSI into the NCEP Global Data Assimilation System. *Weather and Forecasting*, 24, 1691–1705.

Lin, Y. and Mitchell, K. (2005) The NCEP Stage II/IV hourly precipitation analyses: development and applications. *Preprint 19th Conference on Hydrology*. San Diego, CA: American Meteorological Society.

Mansell, E.R., Ziegler, C. and Bruning, E. (2010) Simulated electrification of a small thunderstorm with two-moment bulk microphysics. *Journal of Atmospheric Science*, 67, 171–194. https://doi.org/10.1175/2009JAS2965.1.

Roberts, N.M. and Lean, H.W. (2008) Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Monthly Weather Review*, 136, 78–97.

Schmidt, J., Anderson, A. and Paul, J. (2007) Spatially-variable, physically-derived, flash flood guidance. *Preprints, 21st Conference on Hydrology, San Antonio, Texas*. American Meteorological Society, 6B.2. Available at: https://ams.confex.com/ams/pdfpapers/120022.pdf [Accessed 15th September 2018].

Schumacher, R.S. (2017) Heavy Rainfall and Flash Flooding. *Oxford Research Encyclopedia of Natural Hazard Science.* Oxford, UK: Oxford University Press https://doi.org/10.1093/acrefore/9780199389407.013.132.

Schwartz, C.S., Kain, J.S., Weiss, S.J., Xue, M., Bright, D.R., Kong, F., Thomas, K.W., Levit, J.J., Coniglio, M.C. and Wandishin,

M.S. (2010) Toward improved convection-allowing ensembles: model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Weather and Forecasting*, 25, 263–280.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X.-Y., Wang, W. and Powers, J.G. (2008) *A Description of the Advanced Research WRF Version 3*. NCAR Technical Note NCAR/TN-475+STR, 113 pp.

Skinner, P.S., Wicker, L.J., Wheatley, D.M. and Knopfmeier, K.H. (2016) Application of two spatial verification methods to ensemble forecasts of low-level rotation. *Weather and Forecasting*, 31, 713–735.

Snook, N., Xue, M. and Jung, Y. (2015) Multiscale EnKF assimilation of radar and conventional observations and ensemble forecasting for a tornadic mesoscale convective system. *Monthly Weather Review*, 143, 1035–1057.

Stensrud, D.J., Bao, J.-W. and Warner, T.T. (2000) Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Monthly Weather Review*, 128, 2077–2107.

Stensrud, D.J., Wicker, L.J., Xue, M., Dawson, D.T., II, Yussouf, N., Wheatley, D.M., Thompson, T.E., Snook, N.A., Smith, T.M., Schenkman, A.D., Potvin, C.K., Mansell, E.R., Lei, T., Kuhlman, K.M., Jung, Y., Jones, T.A., Gao, J., Coniglio, M.C., Brooks, H.E. and Brewster, K.A. (2013) Progress and challenges with warn-on-forecast. *Atmospheric Research*, 123, 2–16.

Stensrud, D.J., Xue, M., Wicker, L.J., Kelleher, K.E., Foster, M.P., Schaefer, J.T., Schneider, R.S., Benjamin, S.G., Weygandt, S.S., Ferree, J.T. and Tuell, J.P. (2009) Convective-scale warn-on-forecast system: a vision for 2020. *Bulletin of the American Meteorological Society*, 90, 1487–1499.

Sun, J., Xue, M., Wilson, J.W., Zawadzki, I., Ballard, S.P., Onvlee-Hooimeyer, J., Joe, P., Barker, D.M., Li, P.-W., Golding, B., Xu, M., and Pinto, J. (2014) Use of NWP for Nowcasting Convective precipitation: recent progress and challenges. *Bulletin of the American Meteorological Society*, 95, 409–426.

Toth, Z., Zhu, Y. and Wobus, R. (2004) *March 2004 upgrades of the NCEP global ensemble forecast system*. Available at: http://www.emc.ncep.noaa.gov/gmb/ens/ens_imp_news.html [Accessed 15th September 2018].

Wang, X., Parrish, D., Kleist, D. and Whitaker, J. (2013) GSI 3DVar-based ensemble–variational hybrid data assimilation for NCEP Global Forecast System: single-resolution experiments. *Monthly Weather Review*, 141, 4098–4117. https://doi.org/10.1175/MWR-D-12-00141.1.

Wang, Y. and Wang, X. (2017) Direct assimilation of radar reflectivity without tangent linear and adjoint of the nonlinear observation operator in the GSI-based EnVar system: methodology and experiment with the 8 May 2003 Oklahoma City tornadic supercell. *Monthly Weather Review*, 145, 1447–1471 https://doi.org/10.1175/MWR-D-16-0231.1.

Wei, M., Toth, Z., Wobus, R. and Zhu, Y. (2008) Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus*, 60A, 62.

Wheatley, D.M., Knopfmeier, K.H., Jones, T.A. and Creager, G.J. (2015) Storm-scale data assimilation and ensemble forecasting with the NSSL experimental Warn-on-Forecast System. Part I: radar data experiments. *Weather and Forecasting*, 30, 1795–1817.

Wheatley, D.M., Yussouf, N. and Stensrud, D.J. (2014) Ensemble Kalman filter analyses and forecasts of a severe mesoscale convective system using different choices of microphysics schemes. *Monthly Weather Review*, 142, 3243–3263. https://doi.org/10.1175/MWR-D-13-00260.1.

Whitaker, J.S. and Hamill, T.M. (2012) Evaluating methods to account for system errors in ensemble data assimilation. *Monthly Weather Review*, 140, 3078–3089. https://doi.org/10.1175/MWR-D-11-00276.1.

Whitaker, J.S., Hamill, T.M., Wei, X., Song, Y. and Toth, Z. (2008) Ensemble data assimilation with the NCEP Global Forecast System. *Monthly Weather Review*, 136, 463–482. https://doi.org/10.1175/2007MWR2018.1.

Yussouf, N., Mansell, E.R., Wicker, L.J., Wheatley, D.M. and Stensrud, D.J. (2013a) The ensemble Kalman filter analyses and forecasts of the 8 May 2003 Oklahoma City tornadic supercell storm using single and double moment microphysics schemes. *Monthly Weather Review*, 141, 3388–3412.

Yussouf, N., Gao, J., Stensrud, D.J. and Ge, G. (2013b) The impact of mesoscale environmental uncertainty on the prediction of a tornadic supercell storm using ensemble data assimilation approach. *Advances in Meteorology*, 2013, 1–15.

Yussouf, N., Dowell, D.C., Wicker, L.J., Knopfmeier, K. and Wheatley, D.M. (2015) Storm-scale data assimilation and ensemble forecasts for the 27 April 2011 severe weather outbreak in Alabama. *Monthly Weather Review*, 143, 3044–3066.

Yussouf, N., Kain, J.S. and Clark, A.J. (2016) Short-term probabilistic forecasts of the 31 May 2013 Oklahoma Tornado and flash flood event using a continuous-update-cycle storm-scale ensemble system. *Weather and Forecasting*, 31, 957–983.

How to cite this article: Yussouf N, Knopfmeier KH. Application of the Warn-on-Forecast system for flash-flood-producing heavy convective rainfall events. *Q J R Meteorol Soc*. 2019;145:2385–2403. https://doi.org/10.1002/qj.3568